Amendments to the Claims:

The following listing of claims will replace all prior versions, and listings, of claims in the application:

1. (Currently Amended) A method for improving appearance of captured bilevel image data, comprising:

receiving a degraded input bilevel image;

connecting dark pixels in the input image to adjacent dark pixels to form connected components comprising a set of dark pixels connected either diagonally or orthogonally and surrounded by white pixels;

performing initial clustering of individual connected components;

determining a "most likely" cluster representative by use of a probabilistic model of the scanner used for scanning; and

assembling the sets by substituting the "most-likely" cluster representative for each family member of each cluster set to form an output image,

wherein the step of performing initial clustering includes pair-wise matching connected components and determining a match if the pair are within a certain threshold of matching to form cluster sets, each with one or more family members formed of individual connected components,

the step of determining a "most likely" cluster representative for each cluster set uses a optimization procedure in which at least one iteration of pixel flipping is performed,

multiple iterations of pixel flipping are performed, with a first iteration of pixel flipping being performed with a first threshold, and

a subsequent iteration being performed with a more progressive threshold being applied until a maximized probability is attained.

2. (Canceled)

3. (Currently Amended) The method of claim ~~2~~1, wherein the optimization procedure is a hill-climbing optimization procedure.

4. (Currently Amended) The method of claim ~~2~~1, wherein an initial representative is determined by finding a translation τ of each family member that maximizes the probability that the family member is a given original image to obtain a higher resolution histogram.

5. (Currently Amended) The method of claim ~~2~~1, wherein an initial representative is determined by summing scans of each maximized family member to form a double-resolution histogram.

6. (Canceled).

7. (Original) The method of claim 1, further comprising a step of reclustering by comparing cluster representatives of clusters.

8. (Original) The method of claim 7, wherein the step of reclustering reclusters the cluster sets by comparing cluster representatives of smaller cluster sets with cluster representatives of larger cluster sets and merging the smaller cluster set into the larger cluster set when a normalized probability exceeds a default threshold.

9. (Currently Amended) The method of claim ~~2~~1, wherein the initial clustering includes:

forming a bounding box around each connected component A and B;

aligning connected components A and B to each other by aligning centers of their bounding boxes; and

determining a match of the connected components A and B if:

$$|A| - |A \cap \overline{B}| \le f(|\partial A|) \text{ and } |\mathrm{B}| - |\mathrm{B} \cap \overline{\mathrm{A}}| \le f(|\partial \mathrm{B}|)$$

where:

$|A|$ denotes the number of black pixels in A;

$A \cap B$ denotes the pixels that are black in both A and B;

$\overline{A}$ denotes a one-pixel dilation of the black pixels in A;

∂A denotes the boundary of A, that is, the set of black pixels with white neighbors; and

$f(n)$ equals 0 for $n \leq 3$, and .025n for $n \geq 7$, and interpolates between these two lines for $3 < n < 7$.

10. (Currently Amended) The method of claim 9, ~~further comprising~~wherein stopping of a match of one of or $A \backslash \overline{B}$ or $B \backslash \overline{A}$ includes a set of more than three pixels that can be enclosed by a 3x3 box.

11. (Original) The method of claim 9, wherein the probability that a pixel in row *i* and column *j* has value A_{ij} (black or white) given B and τ is determined as:

$$P[A_{ij}|B,\tau] = \begin{cases} p(w_{ij}(\tau)) & \textit{if } A_{ij} \textit{ is black;} \\ 1 - p(w_{ij}(\tau)) & \textit{if } A_{ij} \textit{ is white.} \end{cases},$$

where τ represents a translation of a sensor grid used to capture the input image with respect to a given original image region B;

$w_{ij}(\tau)$ denotes the weight of black in the given original image region B seen by the sensor grid in row *i* and column *j* based on a point spread function of the sensor grid; and

$p(w_{ij}(\tau))$ denotes a determined probability that the sensor grid's output pixel would be black.

12. (Original) The method of claim 11, wherein individual pixel probabilities within a connected component A are multiplied to obtain a probability $P[A \mid B,\tau]$ that the connected component A is a capture of the given original image region B at translation τ as:

$$P[A|B,\tau] = \prod_{ij} P[A_{ij}|B,\tau]$$

13. (Original) The method of claim 12, wherein τ is optimized over all possible translations as

$$P[A|B] = \max_{\tau} P[A|B,\tau].$$

14. (OriginalThe method of claim 13, wherein connected component A and given original image region B are prealigned by the centroids of their respective bounding boxes.

15. (Original) The method of claim 14, wherein optimization of τ is limited to the nine shortest vectors in the lattice of the bounding box.

16. (Original) The method of claim 13, wherein the probability of an entire cluster set is computed by multiplying the probabilities of each individual family member using

$$P[C|B] = \prod_{A \in C} P[A|B]$$

and the initial "most likely" cluster representative is the one that maximizes P[C|B].

17. (Original) The method of claim 1, further comprising using chain codes to define a priori probabilities to find the cluster representative.

18. (Original) The method of claim 17, wherein the *a priori* probability is computed by determining the product of transition probabilities around all connected components of the cluster representative to attain a value B^i and the "most likely" representative is the B^i with a maximum $P[C|B^i]$.

19. (Original) The method of claim 7, wherein the step of reclustering is performed by normalizing probabilities using

$$N[A_i|B_j] = (P[A_i | B_j])^{1/p},$$

where p is the number of pixels in the connected component A_i (aligned with the representative image B_j) that are within a sensor disk's radius of a black pixel in either the connected component A_i or the representative image B_j.

20. (Currently Amended) The method of claim 19, wherein ~~the~~a threshold is 0.70.

21. (Currently Amended) The method of claim 19, wherein ~~the~~a threshold is 0.68 when ~~the~~a smaller cluster set is a singleton and ~~the~~a larger cluster set has at least four family members.

22. (Currently Amended). The method of claim 7, wherein reclustering stops when the larger cluster sets ~~has~~have fewer than four family members.

23. (Original) The method of claim 1, further comprising a step of breaking run-through letters by computing a sequence of breakable positions of singleton cluster representatives and comparing each breakable position portion with other cluster representatives.

24. (Original) The method of claim 23, further comprising a step of merging a breakable position portion with a cluster set when the comparison indicates a sufficient match.

25. (Original) The method of claim 1, wherein the step of assembly includes aligning centers of bounding boxes and testing double-resolution translations to recompute alignment and determine the most likely position of the cluster representative.

26. (New) A method for improving appearance of captured bilevel image data, comprising:

receiving a degraded input bilevel image;

connecting dark pixels in the input image to adjacent dark pixels to form connected components comprising a set of dark pixels connected either diagonally or orthogonally and

surrounded by white pixels;

performing initial clustering of individual connected components;

determining a "most likely" cluster representative by use of a probabilistic model of the scanner used for scanning;

assembling the sets by substituting the "most-likely" cluster representative for each family member of each cluster set to form an output image; and

reclustering by comparing cluster representatives of clusters, wherein

reclustering stops when the larger cluster sets have fewer than four family members.

27. (New) The method of claim 26, wherein the step of reclustering reclusters the cluster sets by comparing cluster representatives of smaller cluster sets with cluster representatives of larger cluster sets and merging the smaller cluster set into the larger cluster set when a normalized probability exceeds a default threshold.

28. (New) The method of claim 26, wherein the step of reclustering is performed by normalizing probabilities using

$$N[A_i | B_j] = (P[A_i | B_j])^{1/p},$$

where p is the number of pixels in the connected component A_i (aligned with the representative image B_j) that are within a sensor disk's radius of a black pixel in either the connected component A_i or the representative image B_j.

29. (New) The method of claim 28, wherein a threshold is 0.70.

30. (New) The method of claim 28, wherein a threshold is 0.68 when a smaller cluster set is a singleton and a larger cluster set has at least four family members.

31. (New) The method of claim 26, further comprising a step of breaking run-through letters by computing a sequence of breakable positions of singleton cluster

representatives and comparing each breakable position portion with other cluster representatives.

32. (New) The method of claim 31, further comprising a step of merging a breakable position portion with a cluster set when the comparison indicates a sufficient match.